PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 : A61K 37/02, 39/04

A1

(11) International Publication Number: WO 89/12455

(43) International Publication Date: 28 December 1989 (28.12.89)

(21) International Application Number: PCT/US89/02619

(22) International Filing Date: 15 June 1989 (15.06.89)

(30) Priority data:
207,298 15 June 1988 (15.06.88) US

(71) Applicants: WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH [US/US]; Nine Cambridge Center, Cambridge, MA 02142 (US). MEDICAL RESEARCH COUNCIL [GB/GB]; 20 Mount Pleasant, London WIN 4AL (GB).

(72) Inventors: YOUNG, Richard, A. ; 5 Sawmill Brook Road, Winchester, MA 01890 (US). YOUNG, Douglas ; 44 Lawnclose Ruislip, Middlesex HA4 6ED (GB).

(74) Agents: GRANAHAN, Patricia et al.; Hamilton, Brook, Smith & Reynolds, Two Militia Drive, Lexington, MA 02173 (US).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: STRESS PROTEINS AND USES THEREFOR

(57) Abstract

Stress proteins and their use to immunize an individual against a nonviral infection or to induce immune tolerance in an individual, as well as a method of immunizing an individual by administering a selected stress protein and a method of inducing immune tolerance in an individual by administering a selected stress protein.

STRESS PROTEINS AND USES THEREFOR

Description

Background of the Invention

Although the function of stress proteins is not entirely clear, it appears that some participate in assembly and structural stabilization of certain cellular and viral proteins, and their presence at high concentrations may have an additional stabilizing effect during exposure to adverse conditions. Neidhardt, F.C. and R.A. VanBogelen, In: Escherichia coli and Salmonella typhimurium, Cellular and Molecular Biology, (eds. Neidhardt, F.C., Ingraham, J.L., Low, K.B., Magasanik, B. Schaechter, M. and Umbarger, H.E. (Am. Soc. Microbiol., Washington, D.C.), pp. 1334-1345 (1987); Pelham, H.R.B. Cell, 46:959-961 (1986); Takano, T. and T. Kakefuda, Nature, 239:34-37 (1972); Georgopoulos, C. et al., New Biology, 239:38-41 (1972). Phagocytic host cells produce a hostile environment for foreign organisms, and the ability to produce stress proteins has been implicated in the survival of bacterial pathogens within macrophages Christman, M.F. et al., Cell, 41:753-762 (1985).

Mycobacterium (M.) tuberculosis and Mycobacterium (M.) leprae are the etiologic agents of tuberculosis and leprosy, respectively. These diseases afflict 20-30 million people and continue to present a significant global health problem. Joint International Union Against Tuberculosis and World Health Organization Study Group, Tubercle, 63:157-169 (1982); Bloom, B. and T. Godal, Rev. Infect Dis. 5:765-780 (1983). To develop more effective

that most of the antigens identified with monoclonal antibodies are involved in the T-cell response to mycobacterial infection or vaccination in mice and in humans. Limiting dilution analysis indicates that 20% of the mycobacterial-reactive $CD4^+$ T lymphocytes in mice immunized with M. tuberculosis recognize a single protein, the 65-kDa antigen. Kaufman, S.H.E. et al., Eur J. Immunol., 17:351-357 (1987).

Summary of the Invention

The present invention relates to stress proteins and methods of modulating an individual's immune response, either to a pathogen or to his or her own cells, such as occurs in autoimmune diseases. In particular, it relates to the use of such stress proteins as a "vaccine" in immune prophylaxis therapy, which results in an induction or enhancement of immune response to a selected pathogen and as an immunotherapeutic agent in treatment of autoimmune diseases, which results in a decrease of an individual's response to his or her own cells. In immune prophylaxis, stress proteins are administered to prevent or reduce the effects in an individual of a pathogen, which can be any virus, microorganism or other organism or substance (e.g., a toxin or toxoid) which causes disease. In preventing or reducing adverse effects of nonviral pathogens (e.g., bacteria, mycobacterial) according to the method of the present invention, an individual's immune response to the nonviral pathogen's stress protein(s) is induced or enhanced through the administration of a vaccine which includes the pathogen's stress protein(s) and, generally, an adjuvant.

protein can be administered in multiple doses over time in order to induce immune tolerance against an autoimmune disease such as rheumatoid arthritis.

Brief Description of the Drawings

Figure 1 is a graphic representation of the homologies between mycobacterial antigens and known stress proteins. Figure 1A is a representation of sequence similarity between portions of the M. tuberculosis 71-kDa antigen (residues 1-204; TB 71kDa) and the E. coli DnaK protein (residues 430-469). Figure 1B is a representation of sequence similarity between portions of the M. tuberculosis 65-kDa antigen (residues 1-540; TB 65kDa) and the E. coli GroEL protein (residues 1-547).

Figure 2 is a comparison of the amino acid sequence of the human P1 protein (573 residues) and the amino acid sequence of the groEL protein (547 residues).

Figure 3 is a comparison of the amino acid sequence of the human P1 protein (573 residues), which is a homolog of groEL protein, and the amino acid sequence of the 65kDa M. leprae protein (540 residues).

Figure 4 is a comparison of the amino acid sequence of the human P1 protein (573 residues), which is a homolog of the groEL protein, and the amino acid sequence of the 65kDa M. tuberculosis protein (540 residues).

Detailed Description of the Invention

The present invention is based on the observation that stress proteins are among the major antigens available for presentation to T lymphocytes and may be common immune targets in a broad spectrum of infectious

Exp. Med., 165:1430-1435 (1987)) and the malarial parasite Brugia malayi (Selkirk, M.E. et al., J. Cell Biochem., 12D:290 (1988)). Similarly, homologues of GroEL have been found among antigens involved in the immune response to Salmonella typhimurium and Coxiella. Vodkin, M.H. and J.C. Williams, J. Bacteriol, 170:1227 (1988). The presence of stress proteins among major immune targets in a variety of human pathogens is support for the idea that the stress response may be a general component of infection and that stress proteins should be considered among candidates for subunit vaccines. All organisms respond to heat by inducing synthesis of heat shock proteins (hsps), which are a group of proteins. This response is the most highly conserved genetic system known and has been shown to occur in every organism, including microorganisms, plants and animals, investigated to date. Many of the characteristics of the response are common to all organisms and the hsps are among the most highly conserved proteins known. For example, hsp90 family and hsp70 family proteins are present in widely diverse organisms. The proteins in each family--even in such diverse organisms--show approximately 50% identity at the amino acid level and at the nonidentical residues, exhibit many similarities. Several of the proteins induced by heat are also induced by a variety of other stresses. The hsps or a closely related/similar protein are present in all organisms at normal temperatures and have been shown to have key functions in normal cell metabolism. Lindquist, S. and E.A. Craig, Ann. Rev. Genet., 22:631-677 (1988). Because the stress response is common to prokaryotes and

In view of the involvement of proteins of M. tuberculosis and M. leprae in humoral and cell-mediated immune responses and to establish the functions of the these proteins in the mycobacterial cell, the DNA encoding several of the M. tuberculosis and M. leprae antigens have been sequenced. It has been demonstrated, as a result, that many of these mycobacterial protein antigens exhibit striking sequence similarity to known stress-induced proteins. Three of the M. leprae and two of the M. tuberculosis protein antigens studied have been shown to exhibit striking sequence similarity to known stress proteins. For reasons discussed in the Exemplification, it is concluded that two of the M. leprae and two of the M. tuberculosis antigens are homologues of the E. coli DnaK and GroEL proteins.

In experimental mice, immunization with mycobacterial lysates elicits antibody responses to at least six M. tuberculosis protein antigens and a similar number of M. leprae protein antigens. Monoclonal antibodies specific for these proteins have been used to isolate clones from λgt11 DNA expression libraries of M. tuberculosis and M. leprae. The sequence of the DNA clones revealed that mycobacterial hsp70 (alias 70 kDa antigen) and hsp60 (alias 65 kDa antigen, groEL) were the major targets of the murine antibody response to both M. tuberculosis and M. leprae. Two additional hsp's, an 18 kDa member of the small hsp family and a 12 kDa homologue of groES, were found among the M. leprae and M. tuberculosis antigens. Young, D.B., et al., Proc. Natl. Acad. Sci. USA, 85:4267-4270 (1988); Shinnick, T.M., et al., Nuc. Acids Res., 17:1254 (1989).

some fraction of the human T cell clones were shown to recognize one or more of the mycobacterial stress proteins.

Stress Proteins are Immune Targets in Non-viral Infections

The observation that stress proteins are important targets of the immune response to mycobacterial infection and the knowledge that the major stress proteins are conserved and abundant in other organisms suggested that stress proteins are likely to be immune targets in many non-viral infections. Indeed, that is now clearly the case. Antigens from a wide variety of infectious agents have been identified as members of stress protein families. The major stress protein antigen recognized by antibodies in bacterial infections is hsp60. "Common antigen", an immunodominant protein antigen long known to be shared by most bacterial species, turns out to be hsp60. Shinnick, T.M., et al., Infect. Immun., 56:446 (1988); Thole, J.E.R., et al., Microbial Pathogenesis, 4:71-83 (1988). Stress proteins have also been identified as immune targets in most major human parasite infections. Bianco, A.E., et al., Proc. Natl. Acad. Sci., USA, 83;8713 (1986); Nene, V., et al., Mol. Biochem. Parasitol., 21:179 (1986); Ardeshir, F., et al., EMBO J., 6:493 (1987); Hedstrom, R., et al., J. Exp. Med., 165:1430 (1987); Selkirk, M.E., et al., J. Cell Biochem., 12D:290 (1988); Engman, D.M., et al., J. Cell Biochem., 12D: Supplement, 290 (1988); Smith, D.F., et al., J. Cell Biochem., 12D:296 (1988). Antibodies to hsp70 have been identified in the sera of patients

of years. A number of different proteins present in synovial membranes have been proposed to be the cross-reactive rat antigen, but were later discounted as procedures for the purification of these proteins improved. van Eden, W., et al., Proc. Natl. Acad. Sci., USA, 82:5117-5120 (1985); Holoshitz, J., et al., Science, 219:56-58 (1983). The M. tuberculosis antigen recognized by the arthritogenic T cells was recently shown to be a 65 kDa protein (van Eden, W., et al., Nature, 331:171 (1988), which has now been shown to be hsp60 (see the Exemplification). Using a combination of truncated recombinant 65 kDa proteins and peptides, a nine amino acid epitope of hsp60 has been identified as the minimum stimulatory sequence for arthritogenic T cell clones in proliferation assays. Now that it is clear that some arthritogenic T cells recognize the mycobacterial hsp60, it is quite possible that the rat autoantigen is also hsp60.

The results obtained in the adjuvant arthritis model led investigators to determine whether T lymphocytes from human rheumatoid arthritis patients also recognize mycobacterial antigens. These investigators have found not only that patients with rheumatoid arthritis have T cells that recognize M. tuberculosis antigens, but that these T cells have diverse phenotypes. Substantial proliferative responses to mycobacterial extracts are observed with uncloned T cells (predominantly $CD4^+$) from both synovial infiltrates and peripheral blood, although responses are generally greater in synovial infiltrates. Abrahamson, T.G., et al., Scand. J. Immunol., 7:81-90 (1978); Holoshitz, J., et al., Lancet ii, 305-306 (1986).

self stress protein determinants; and observations that stress responses are induced by viral infection and by cell transformation, all suggest a model of immune surveillance in which self-reactive T cells provide a first line of defense against infection and transformation by recognizing and helping to eliminate stressed autologous cells, as well as cells infected with intracellular bacteria. The pool of lymphocytes that recognize conserved stress protein determinants might be induced during establishment of natural microbial flora on the skin and in the gut, and maintained by frequent stimulation by bacteria and viruses as well as other stressful stimuli encountered during a normal lifetime. This model is attractive because it provides a way in which the immune system could exploit the existence of conserved epitopes in stress proteins to respond immediately to antigenically diverse pathogens and cellular changes, producing an initial defense that need not await the development of immunity to novel antigens.

Stress protein induction occurs in eukaryotic cells following infection by diverse viruses in vitro. Collins, P.L., and Hightower, L.E., J. Virol., 44:703-707 (1982); Nevins, J.R., Cell, 29:913-939 (1982); Garry, R.F., et al., Virology, 129:391-332 (1988); Khandjian, E.W. and Turler, H., Mol. Cell Biol., 3:1-8 (1983); LaThangue, N.B., et al., EMBO J., 3:267-277 (1984). CTL that recognize these neo-antigens could limit the spread of virus by killing infected cells, possibly before substantial amounts of mature virus are assembled, and by secreting the lymphokine γ-interferon. Pestka, S., in: Methods Enzymol., Interferons, Part A., Vol. 79, Academic

proteins are constitutively expressed in normal cells, although at lower levels than in stressed cells, the potential for autoreactivity is ever-present. Normal cells may escape destruction by expressing only substimulatory levels of stress protein determinants on their surfaces. In addition, stress proteins may only be processed and presented during stress, and it may be relevant that many stress proteins have altered intracellular locations during stress. Finally, immune regulatory networks may prevent activation of autoreactive T cells under normal conditions. The regulatory constraints required by this system might occasionally break down, perhaps during stress caused by bacterial or viral infections, leading to autoimmune disease. Rhematoid arthritis may be such a disease.

Modulation of Immune Response

The precise relationship between stress proteins and the host immune response to infection is as yet undefined. When cells are subjected to a variety of stresses, they respond by selectively increasing the synthesis of a limited set of stress proteins. Some stress proteins, including the products of dnaK and groEL, are major constituents of the cell under normal growth conditions and are induced to even higher levels during stress. Lindquist, S., Annu. Rev. Biochem., 55: 1151-1191 (1986); Neidhardt, F.C. and R.A. VanBogelen, In Escherichia coli and Salmonella Typhimurium, Cellular and Molecular Biology, (eds. Neidhardt, F.C., Ingraham, J.L. Low, K.B. Magasanik, B. Schaechter, M. and Umbarger, H.E.) Am. Soc. Microbiol., Washington, D.C., pp. 1134-1345 (1987). It has now been demonstrated that stress-

response (and, thus, reducing the pathogen's effects) can be used.

First, because the nonviral pathogen's stress proteins are distinguishable from those of the host, it is possible to induce an immunoprophylactic response specific to the pathogen's stress proteins. This can be carried out by administering a vaccine which includes all or a portion (e.g., sufficient sequence to have the desired stimulatory effect on immune response) of the pathogen's stress protein or of another protein having an amino acid sequence sufficiently similar to that of the stress protein sequence to stimulate the immune response to the pathogen. Alternatively, highly conserved stress protein determinants, such as those shown to be recognized by a variety of T cells, can be administered as a type of "general" vaccine. In either case, the immune response to the stress protein sequence will be increased and effects of the nonviral pathogen will be reduced (decreased, prevented or eliminated).

Second, it is also possible to induce or enhance the immune surveillance system or response which is directed to stressed host cells. This is described further in the context of enhancing immune response in those instances in which the pathogen (e.g., a virus, transforming agent) does not have (express) its own stress proteins (i.e., stress proteins distinguishable from host stress proteins).

The vaccine administered to induce or enhance immune response to nonviral pathogens includes a stress protein of the pathogen against which an immune response is desired, a portion of that protein of sufficient size to

or any other substances or changes in condition which induce the stress response in the individual being treated. (This can also be employed in conjunction with the vaccine, described previously, administered to enhance immune response to a stress protein-producing pathogen.) It is known that increased levels of stress proteins are produced in many types of cancer cells. Enhancement of the immune surveillance system, as described, can be used to facilitate destruction and/or to prevent progression or establishment of cancer cells.

The method of the present invention can also be used to modify or modulate an individual's response to his or her own cells (e.g., as in autoimmune diseases). There are at least two ways in which the present invention can be used immunotherapeutically. First, stress proteins, such as heat shock protein (hsp) 70 and hsp60, are known to be involved in autoimmune disease. It is, thus, possible to turn down an individual's response to "self" by administering the appropriate stress protein(s) in such a manner that the individual becomes more tolerant of the protein. Second, because it is known that the immune response in autoimmune diseases is to stress proteins, it is possible to selectively inhibit or interfere with the ability of immune cells which normally interact with such proteins to do so. This can be done, for example, by administering monoclonal antibodies that bind to specific T cell receptors and delete or disable such cells. Alternatively, rather than knocking out immune cells, the stress response in all cells can be turned down by administering a drug capable of reducing a cell's ability to undergo the stress response. For

(1980). DNA sequences were determined for both strands of DNA. Computer analysis of sequences with UWGCG programs was as described by Devereux, J., et al. Nucleic Acids Res., 12:387-395 (1984).

Immunoblot Analysis. Escherichia coil strain TG1 was transformed with the following plasmids by standard procedures (Maniatis, T., et al., Molecular Cloning, A Laboratory Manual (Cold Spring Harbor Lab., Cold Spring Harbor, NY) (1982), with selection for ampicllin resistance: pND5, a derivatrive of pBR325 containing the E. coli groE genes (Jenkins, A.J., et al., Mol. Gen. Genet. 202:446-454 (1986); pUC8 (Vic?, J., Gene, 19:259-268 (1982); pUC8 with insert DNA for λgt11 clone Y3178 (M. leprae 65-kDa antigen, Young, R.A., et al., Nature, (London) 316:450-452 (1985)) ligated in the EcoRI site.

Overnight cultures of E. coli strains in Luria-Bertani (LB) medium were centrifuged and resuspended in isotonic phosphate-buffered saline at a cell density corresponding to an absorbance of 20 at 60 nm. An equal volume of sample buffer containing 2% (wt/vol) polycrylamide gels in the presence of $NaDodSO_4$ was added, and, after heating on a boiling water bath for 2 min, 5-ml samples were electrophoresed on 12% (wt/vol) polycrylamide gels in the presence of $NaDodSO_4$. Blots were prepared by electrophoretic transfer of the proteins to a nitrocellulose membrane, and binding of monoclonal antibodies was assayed with a peroxidase-conjugated secondary antibody as described. Young., D.B., et al., Infect. Immun., 55:1421-1425 (1987).

TABLE Mycobacterial protein antigens

Protein, kDA	Recognized by human T cells	Subjected to sequence analysis	Homology with known proteins
M. tuberculosis			
71	+	+	DnaK
65*	+	+	GroEL
38	+	-	-
19	+	+	None
14	+	-	-
12	ND	-	-
M. leprae			
70	ND	-	DnaK
65	+	+	GroEL
36	+	-	-
28	+	-	-
18	+	+	Plant Hsp
12	ND	-	-

Mycobacterial protein antigens, their recognition by human T cells, and homology of the deduced mycobacterial protein sequences to known proteins are summarized.
ND, not determined; +, yes; -, no

* Includes data derived from study of the 65-kDA antigens of M. bovis BCG (bacillus Calmette-Guerin), which is identical to the M. tuberculosis 65-kDA antigen.

+ A. S. Mustafa, J. R. Lamb, D. Young and R A. Young, unpublished data.

The Mycobacterial 65-kDa antigen. The 65-kDa antigens of M. tuberculosis and M. leprae are involved in in the human T-cell response to mycobacterial infection (Table). Genes encoding these proteins have been isolated (Husson, R.N. and Young, R.A., Proc. Natl. Acad. Sci., USA, 84:1679-1683 (1987); Young, R.A., et al., Nature, (London) 316:450-452 (1985)) and sequenced (Shinnick, T.M., J. Bacteriol., 169:1080-1088 (1987); Mehram, V., et al., Proc. Natl. Acad. Sci., USA, 83:7013-7017 (1986)), revealing that the amino acid sequences of the 65-kDa antigens of M. tuberculosis and M. leprae are 95% identical. These protein sequences exhibit no significant sequence similarity to proteins in the GenBank database.

Identification of these proteins was based on the observation that some monoclonal antibodies directed against the mycobacterial 65-kDa antigens cross-react with an E. coli protein of 60kDa. E. coli cells transformed with the plasmid pND5 (Sanger, F., et al., Proc. Natl. Acad. Sci., USA, 74:5463-5467 (1977), which contains the E. coli gro E genes, had been shown to accumulate large amounts of the 60-kDa protein. A comparison of the mycobacterial 65-kDa protein sequences with those determined for E. coli groEL (C. Woolford, K. Tilly, C. Georgopoulous, and R.H., unpublished data) revealed the extent of the sequence similarity as shown in Fig. 1B.

The 60-kDa Gro EL protein is a major stress protein in E. coli. Lindquist, S., Annual Rev. Biochem., 55:1151-1191 (1986); Nature, 333:330-334 (1988). There is some evidence that the mycobacterial 65-kDa proteins

CLAIMS

1. A vaccine comprising all or a portion of a selected stress protein or all or a portion of a protein having an amino acid sequence sufficiently homologous to the amino acid sequence of the stress protein.

2. A vaccine of Claim 1 in which the stress protein is a mycobacterial stress protein or a protein having an amino acid sequence sufficiently homologous to the amino acid sequence of the mycobacterial stress protein.

3. A composition for use as an agent to induce immune tolerance, comprising a selected stress protein.

4. A composition for use in treating an autoimmune disease, comprising all or a portion of a selected stress protein or all or a portion of a protein having an amino acid sequence sufficiently homologous to the amino acid sequence of the stress protein.

5. A composition of Claim 4 for treating rheumatoid arthritis.

6. A vaccine for use in enhancing in an individual the immune response to a nonviral pathogen, comprising all or a portion of a stress protein of the nonviral pathogen against which the enhanced response is desired.

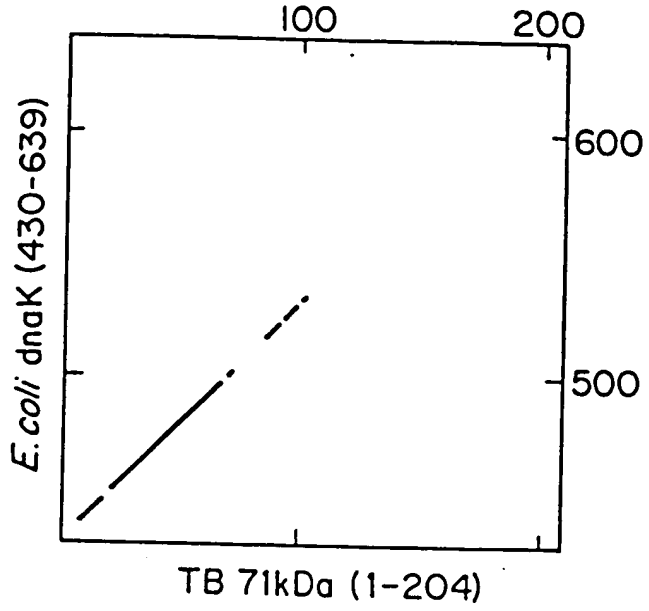

100
200
600
500
E.coli dnaK (430-639)
TB 71kDa (1-204)

Fig. 1A

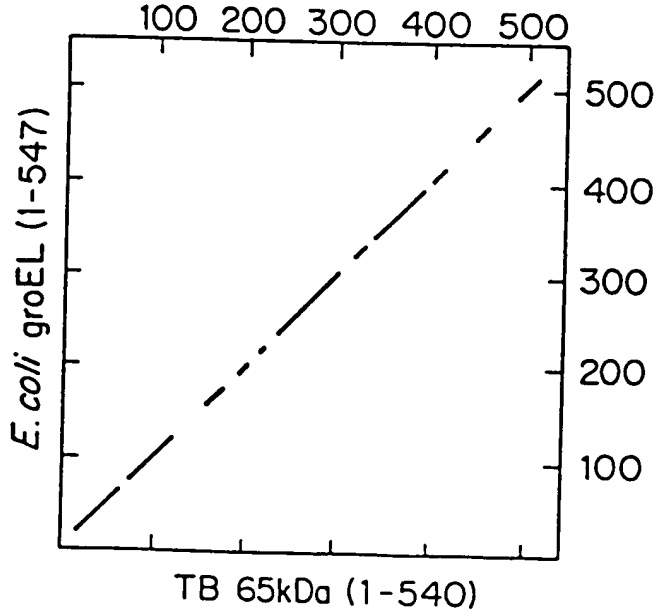
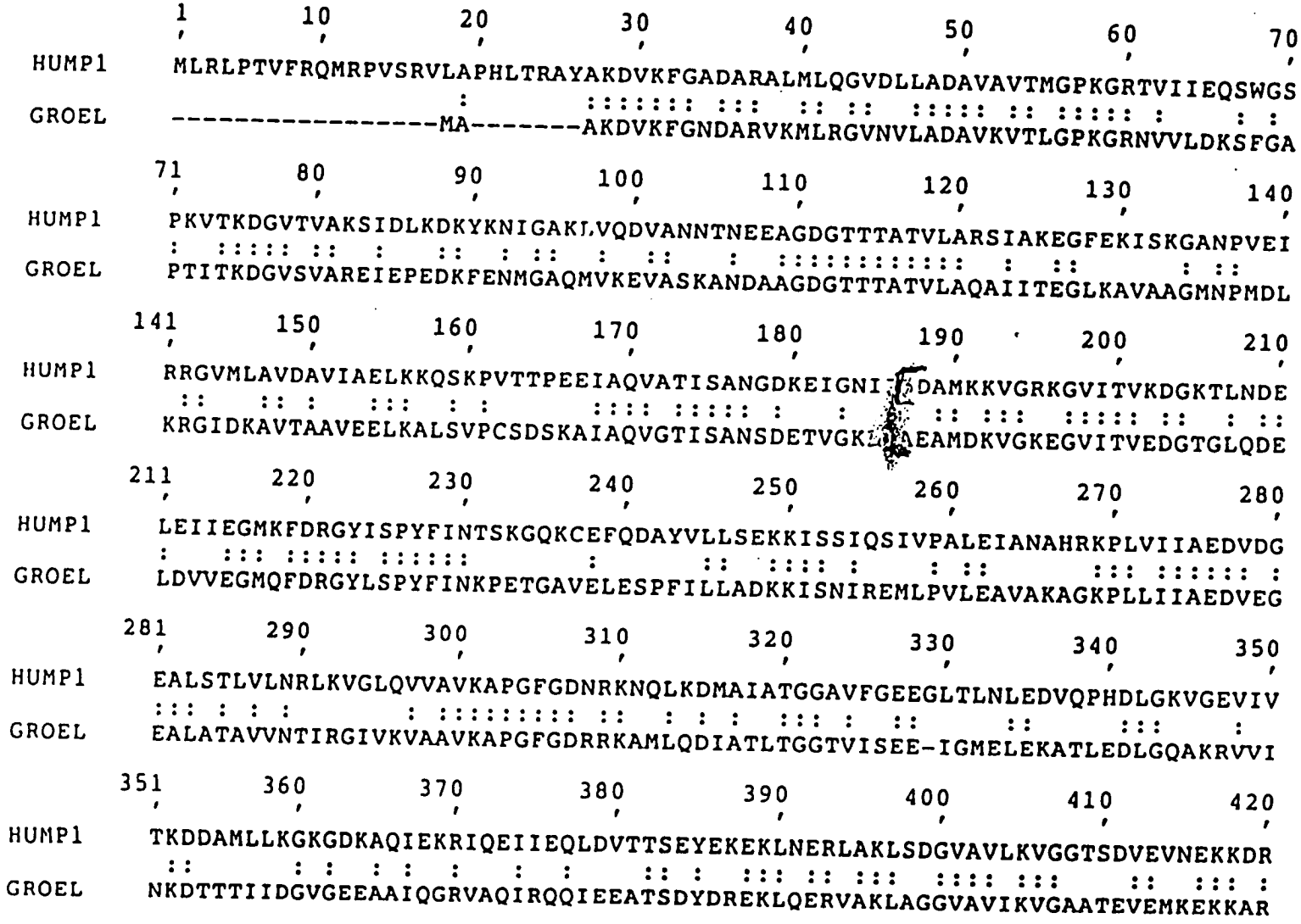

100 200 300 400 500
500
400
300
200
100
E.coli groEL (1-547)
TB 65kDa (1-540)

Fig. 1B

FIGURE 2



SUBSTITUTE SHEET

FIGURE 2 CONT'D

```
        421       430       440       450       460       470       480       490
HUMP1     VTDALNATRAAVEEGIVLGGGCALLRCIPALDSLTPANEDQKIGIEIIKRTLKIPAMTIAKNAGVEGSLI
          : ::: :::::::::: : ::: :: : :   :  :   ::::          :   :  : : : : :
GROEL     VEDALHATRAAVEEGVVAGGGVALIRVASKLADLRGQNEDQNVVSSSL-RAMEAPLRQIVLNCGEEPSVV
        491       500       510       520       530       540       550       560
HUMP1     VEKIMQSSSEVGYDAMAGDFVNMVEKGIIDPTKVVRTALLDAAGVASLLTTAEVVVTEIPKEEKDPGMGA
                  :: :       ::   :: ::::: : ::  :: :: :  : :  ::  ::        ::
GROEL     ANTVKGGDGNYGYNAATEEYGNMIDMGILDPTKVTRSALQYAASVAGLMITTECMVTDLPKND-AADLGA
        561       570
HUMP1     MGGMGG--GMGGGMF
           :::::  :::: :
GROEL     AGGMGGMGGMGGMM-
```

Total score = 4667, 5 breaks
276 identities out of 545 possible matches between residues

25 random runs
Alignment score = 65.34 SD Standard deviation = 18.94 Mean = 3429.48

FIGURE 3

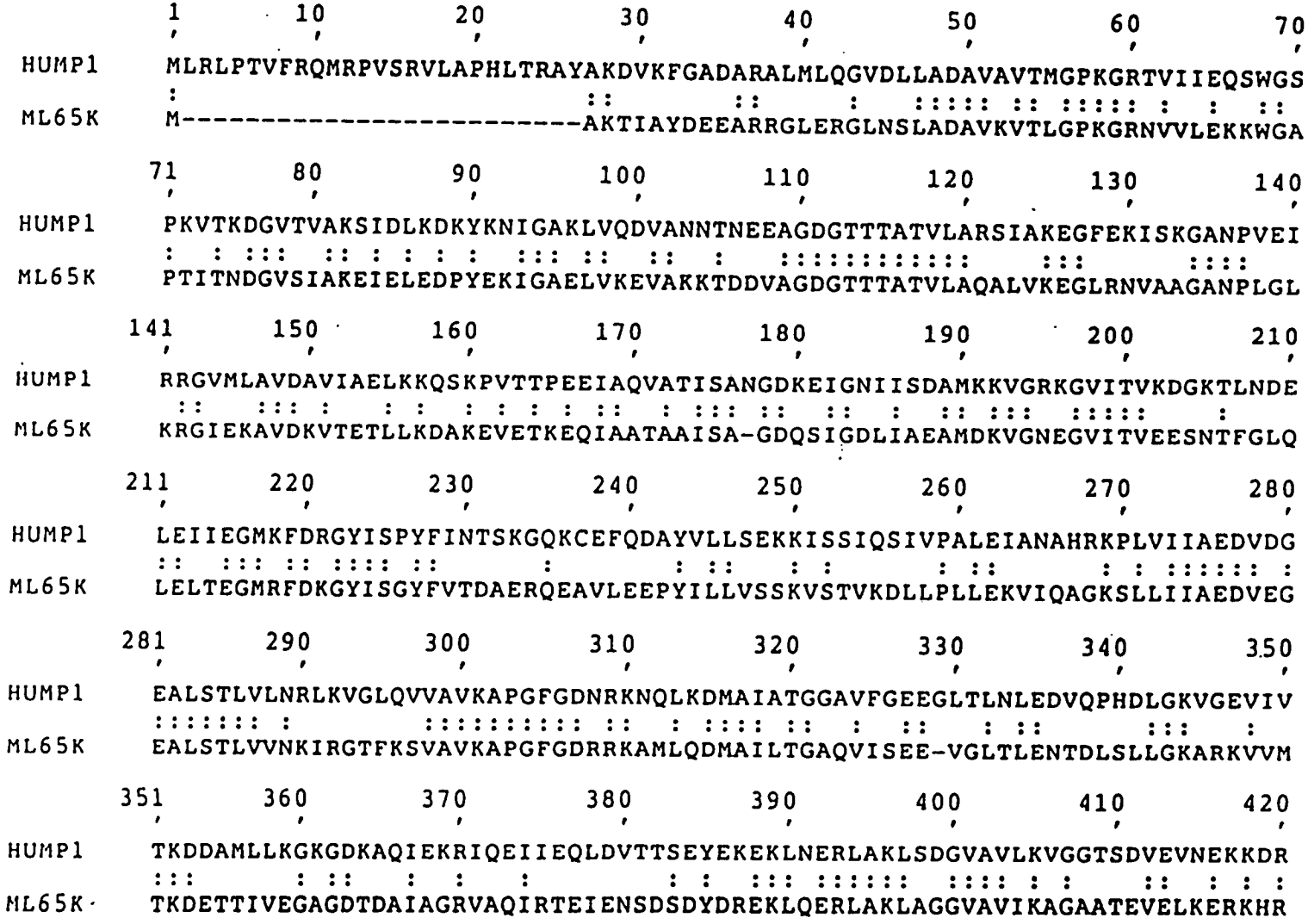


FIGURE 3 CONT'D

```
          421       430       440       450       460       470       480       490
HUMP1     VTDALNATRAAVEEGIVLGGGCALLRCIPALDSLTPANEDQKIGIEIIKRTLKIPAMTIAKNAGVEGSLI
            ::      :::::::: :::  ::   :::: :          :  : : :  :  :: : : :
ML65K     IEDAVRNAKAAVEEGIVAGGGVTLLQAAPALDKLKLTGDEAT-GANIVKVALEAPLKQIAFNSGMEPGVV

          491       500       510       520       530       540       550       560
HUMP1     VEKIMQSSSEVGYDAMAGDFVNMVEKGIIDPTKVVRTALLDAAGVASLLTTAEVVVTEIPKEEKDPGMGA
           ::    :   :  :  :          :  :: :: : ::  ::  : : :: : ::   :     :
ML65K     AEKVRNLSVGHGLNAATGEYEDLLKAGVADPVKVTRSALQNAASIAGLFTT-EAVVADKPEKTAAPASDP

          561       570
HUMP1     MGGMGGGMGGGMF
           ::::: :    :
ML65K     TGGMGG-MD---F
```

Total score = 4552, 7 breaks
255 identities out of 540 possible matches between residues

25 random runs
Alignment score = 47.73 SD Standard deviation = 23.86 Mean = 3413.16

FIGURE 4

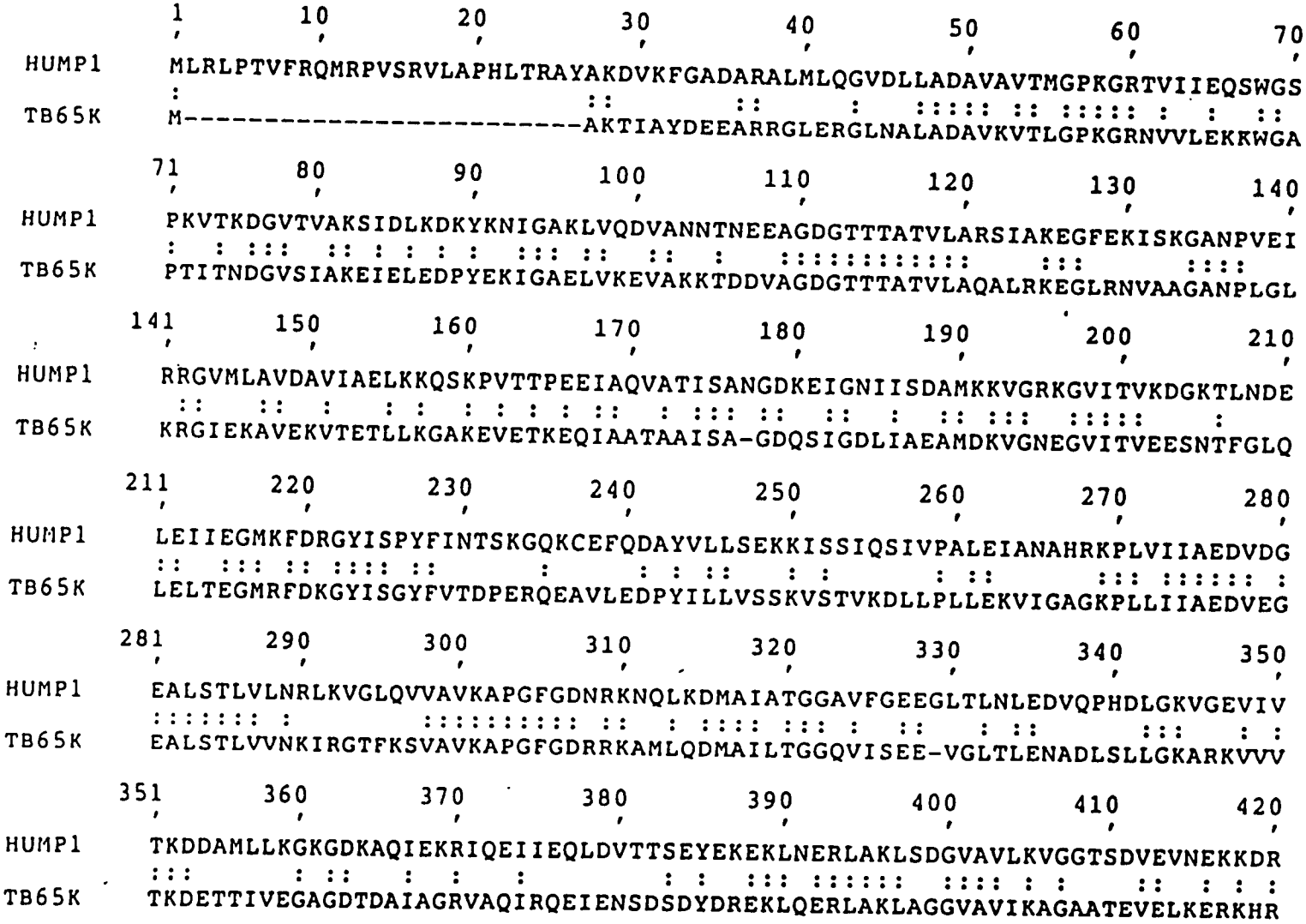

1 10 20 30 40 50 60 70
HUMP1 MLRLPTVFRQMRPVSRVLAPHLTRAYAKDVKFGADARALMLQGVDLLADAVAVTMGPKGRTVIIEQSWGS
TB65K M----------AKTIAYDEEARRGLERGLNALADAVKVTLGPKGRNVVLEKKWGA
71 80 90 100 110 120 130 140
HUMP1 PKVTKDGVTVAKSIDLKDKYKNIGAKLVQDVANNTNEEAGDGTTTATVLARSIAKEGFEKISKGANPVEI
TB65K PTITNDGVSIAKEIELEDPYEKIGAELVKEVAKKTDDVAGDGTTTATVLAQALRKEGLRNVAAGANPLGL
141 150 160 170 180 190 200 210
HUMP1 RRGVMLAVDAVIAELKKQSKPVTTPEEIAQVATISANGDKEIGNIISDAMKKVGRKGVITVKDGKTLNDE
TB65K KRGIEKAVEKVTETLLKGAKEVETKEQIAATAAISA-GDQSIGDLIAEAMDKVGNEGVITVEESNTFGLQ
211 220 230 240 250 260 270 280
HUMP1 LEIIEGMKFDRGYISPYFINTSKGQKCEFQDAYVLLSEKKISSIQSIVPALEIANAHRKPLVIIAEDVDG
TB65K LELTEGMRFDKGYISGYFVTDPERQEAVLEDPYILLVSSKVSTVKDLLPLLEKVIGAGKPLLIIAEDVEG
281 290 300 310 320 330 340 350
HUMP1 EALSTLVLNRLKVGLQVVAVKAPGFGDNRKNQLKDMAIATGGAVFGEEGLTLNLEDVQPHDLGKVGEVIV
TB65K EALSTLVVNKIRGTFKSVAVKAPGFGDRRKAMLQDMAILTGGQVISEE-VGLTLENADLSLLGKARKVVV
351 360 370 380 390 400 410 420
HUMP1 TKDDAMLLKGKGDKAQIEKRIQEIIEQLDVTTSEYEKEKLNERLAKLSDGVAVLKVGGTSDVEVNEKKDR
TB65K TKDETTIVEGAGDTDAIAGRVAQIRQEIENSDSDYDREKLQERLAKLAGGVAVIKAGAATEVELKERKHR

FIGURE 4 CONT'D

```
          421       430       440       450       460       470       480       490
HUMP1     VTDALNATRAAVEEGIVLGGGCALLRCIPALDSLTPANEDQKIGIEIIKRTLKIPAMTIAKNAGVEGSLI
            ::      :::::::: :::   ::    : :: :      :   :  : :  :  :   :: : : :
TB65K     IEDAVRNAKAAVEEGIVAGGGVTLLQAAPTLDELK-LEGDEATGANIVKVALEAPLKQIAFNSGLEPGVV

          491       500       510       520       530       540       550       560
HUMP1     VEKIMQSSSEVGYDAMAGDFVNMVEKGIIDPTKVVRTALLDAAGVASLLTTAEVVVTEIPKEEKDPGMGA
           ::         :  :  :         :  :: :: : ::  ::  :  :  : : ::   :  ::    :
TB65K     AEKVRNLPAGHGLNAQTGVYEDLLAAGVADPVKVTRSALQNAASIAGLFLTTEAVVADKPEKEKASVPG-

          561       570
HUMP1     MGGMGGGMGGGMF
              :: :::  :
TB65K     ----GGDMGGMDF
```

Total score = 4560, 5 breaks
257 identities out of 540 possible matches between residues

25 random runs
Alignment score = 49.36 SD Standard deviation = 23.23 Mean = 3413.16

SUBSTITUTE SHEET